DOCUMENT RESUME

ED 025 762 AL 001 695

By-Park, Nahm-Sheik
Some Basic Assumptions in Transformational Linguistics.
Pub Date 30 Jun 68
Note-21p.
Journal Cit-Language Research; v4 n1 Jun 30 1968
EDRS Price MF-$0.25 HC-$1.15
Descriptors-Child Language, *Deep Structure, Grammar, Linguistic Competence, Linguistic Performance, Linguistic Theory, Structural Analysis, *Surface Structure, Syntax, *Transformation Generative Grammar, *Transformation Theory (Language)

The discussion throughout this paper is devoted to answering the question: What is the nature of our knowledge of language and what theoretical assumptions does the answer entail for linguistic description? Discussed are--(1) what it means to know a language, (2) the distinction between linguistic competence and performance, (3) justification of the transformational model as the theory of linguistic competence, (4) rule-governed creativity versus simple analogy, (5) underlying and superficial structures, (6) transformational grammar and its organization, (7) the theory of language and the grammar of a language, (8) the language acquisition device and the construction of the theory of language, (9) evaluation measure, and (10) justification of grammars. A short bibliography is appended. (DO)

ED025762

Some Basic Assumptions in Transformational Linguistics

Nahm-Sheik Park
(Seoul National University)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Nahm-Sheik Park

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Language Research Vol. IV, No. 1

June 30, 1968

Language Research Institute

Seoul National University

AL 001 695

Some Basic Assumptions in Transformational Linguistics

Nahm-Sheik Park
(Seoul National University)

O. Introduction

The present paper is the first of the four chapters in my unpublished paper "A General Survey of Transformational Model of Linguistic Description".[1] My main aim here is to present some of the fundamental assumptions of transformational linguistics.

The discussion throughout the paper is devoted to answering the question: WHAT IS THE NATURE OF OUR KNOWLEDGE OF LANGUAGE AND WHAT THEORETICAL ASSUMPTIONS DOES THE ANSWER ENTAIL FOR LINGUISTIC DESCRIPTION?[2] No attempt has been made to present formal properties of the rules of grammar that supports the discussion in this paper.[3]

Any possible misrepresentation of the transformational theory of linguistics in this paper is entirely my responsibility. The views represented, however, are due mostly to the group of linguists known as transformationalists. As the title indicates, no claim is made to the comprehensiveness of coverage or, for that matter, up-to-dateness, either.[4]

1. What It Means to Know a Language

The most striking fact about our knowledge of language is that once we acquire and use a language we are able to produce and understand an infinite number of novel sentences coherently and consistently on any relevant occasion, telling grammatical ones from ungrammatical ones, imposing certain interpretations on some ungrammatical ones, etc. That the sentences we know how to deal with after we acquire a language are infinite and novel, i.e., different from one another, can easily be confirmed by 1) considering that there is no longest sentence and

[1] This paper was written in 1966 at the University of Hawaii for a seminar in ethno-linguistics.

[2] For a comprehensive discussion of the answer to this question, the reader is referred to most of Noam Chomsky's works listed in the bibliography.

[3] For details concerning the formal properties of the rules of grammar, see especially Noam Chomsky, "The Notion 'Rule of Grammar'" in J. A. Fodor and J. J. Katz (eds.), *The Structure of Language* (Englewood Cliffs N.J.: Prentice-Hall, 1964).

[4] Transformational theory, especially as it relates to syntax, is in a state of flux.

that there are infinite or almost infinite possibilities of combining "familiar words" in sentences,[5] and 2) reading volumes of books or listening to endless conversations almost never encountering two identical sentences except for greetings, cliches, memorized sentences, and so forth.

According to one estimation, the total number of twenty word sentences in English would be 10^{20} which a native speaker of English tacitly knows, and it would take 100,000,000,000 centuries just to utter them.[6] Consider the amazing productivity represented in 10^{20} sentences built out of 10^4 or so of words. The total number of English sentences of all the varying lengths would run to an infinite astronomical figure, because there is no longest sentence. The same would be the case with all other languages.

It is virtually impossible for human beings with a finite life span and a finite memory to have actually encountered and stored this infinite set of sentences in their memory to be able to produce, understand, and use them so coherently. After all, children acquire language, i.e., the ability to control the use of infinitely many sentences, in the matter of a few years. What is it that makes possible this astounding achievement? From mathematics, we know that our mastery of a finite set of rules enables us to enumerate an infinite set of numbers uniquely. A single rule of the form +integer→1(+integer) enables us to enumerate uniquely the infinite set of positive integers. Our mastery of the finite multiplication table enables us to multiply any arbitrary numbers and get an infinite number of answers. The only plausible explanation for our unbounded ability in the production and interpretation of sentences seems to be our mastery or internalization of a finite set of rules that is capable of enumerating all the sentences of our language in some well-defined way. To know a language is to know this finite set of rules and know in principle how to apply them in producing and understanding infinitely many different sentences.

2. The Distinction between Linguistic Competence and Performance

We have argued that our unbounded ability in language use comes from our mastery of a finite set of rules. However, our knowledge of the rules does not always match our ability to

[5] Since we can make a sentence as long as we desire by using relative modifier clauses, stringing together adjectives modifying nouns, etc., there is no longest sentence, which alone suffices to confirm that the set of sentences that we tacitly know is infinite; apart from the preceding remark, one can also convince himself of the infinitude of novel sentences by considering the amazing fact that given 16 English words "the, this, that, a, good, young, old, boy, girl, man, woman, likes, hates, loves, respects", we can form 7414 sentences (Calculation: mine). As the number of words given increases, the number of possible sentences increases geometrically.

[6] George A. Miller, "The Psycholinguists on the New Scientists of Language", in C.E. Osgood and T.A. Sebeok (eds). *Psycholinguistics: A Survey of Theory and Research Problems* (Bloomington: Indiana University Press, 1965).

apply them in actual speech situations. Although we know the multiplication table perfectly well, we find it dfficult or impossible to multiply immediately $234\times529\times60$, $96,250\times10,601$, etc. This does not mean that we are in principle unable to multiply them. Given paper and pencil plus enough time, we can multiply them and also any arbitrary numbers. This means that although we have the tacit knowledge of the rules, the immediate on-the-spot application of these rules is affected by time limitations, bounds on memory, and other factors that interfere with the application of the rules. By the same token, our knowledge of the system of linguistic rules does not necessarily match our ability to apply them immediately in actual situations. Although our basic knowledge of linguistic rules remains constant, such linguistically irrelevant factors as the organization and limitation of memory, emotional states, external stimulus, educational training, etc, interfere with the on-the-spot application of the rules. As a result, sentences, which are grammatical, i.e., which result from the application of certain rules of grammar, may be too complicated for our immediate comprehension and therefore may not be used in actual speech, due not to the ungrammaticalness of the sentences but to the interference of linguistically extraneous factors such as the ones cited above.

Our tacit knowledge of the system of rules of a language and their application under the idealized condition of freedom from the interference of linguistically irrelevant factors is called our linguistic competence. Transformational model is a theory of linguistic competence and as such is not a model of how speaker-hearers actually go about their business of using sentences.

Our actual performance in the use of language is called technically our "linguistic performance." This linguistic performance, together with introspective reports of native speakers, provides major evidence for linguistic competence. As we shall see, the transformational model as the theory of linguistic competence provides a systematic basis for the study of linguistic performance.

Our tacit knlowledge of the rules of a language tells us that certain sentences are grammatical if they result from the application of the rules only, that others are ungrammatical if they result from the violation of some or all of the rules, that certain sentences are ambiguous, unambiguous, etc.[7] All of these constitute the basis for our linguistic competence and give us the native intuition about our languages.

3. Justification of Transformational Model as the Theory of Linguistic Competence

At present we are almost ignorant of the precise nature of the contribution to linguistic

[7] It is wrong to presume that we are conscious of our knowledge of the rules. However, we can be made conscious of this knowledge.

performance of such factors as emotional states, external stimuli, and organization and limitation of memory.

It thus seems safe to conclude that the feasibility of any linguistic theory that claims to be capable of accounting fully for linguistic performance is seriously doubtful and that its claim is empty or derives from a totally different interpretation of the notion "linguistic performance" from the one we have defined above. In one of his lectures at the Linguistic Institute of summer, 1966, Sydney M. Lamb said that his stratificational theory is superior to transformational theory partly because his theory can account for linguistic performance as well as linguistic competence.[8] It is not clear at all how he is going to account for false starts, speech impediments, etc., which according to our definition fall under linguistic performance. In the absence of his clear-cut supporting evidence, it is hard to understand why he considers that this totally unsupported claim should make his theory superior to transformational model. To the extent that his theory fails to account for what it claims to be capable of accounting for, it is empty and probably wrong.

In view of the above fact, the delimitation of the subject matter of linguistic description to the explication of linguistic competence is well-motivated and justified methodologically and empirically. This is not to deny the importance of the theory of linguistic performance. On the contrary, linguists, psychologists, and neuro-physiologists should cooperate in developing an adequate theory of performance by contributing their shares to such a theory.

This delimitation is further justified by the results of recent psychological experiments, which show that transformational theory as the theory of linguistic competence provides a systematic linguistic basis for the study of linguistic performance. Psychologist Savin, professor at the University of Pennsylvania, found that kernel sentences are easier to remember than non-kernel sentences even when the former is much longer than the latter.[9] Miller and Isard at Harvard found that just one degree of self-embedding does not make a sentence difficult to comprehend, memorize, or use, but that with an increase in the degree of self-embedding the complication

---

[8] Here I am not questioning the adequacy of Lamb's theory. Our concern with linguistic performance is a legitimate one. Transformationalists' position here is that linguistic competence is just one of the factors that contribute to linguistic perfomance and that it is their immediate goal to clarify the nature of linguistic competence.

[9] Using such terms as "kernel" and "non-kernel" even before defining them seems like putting the cart before the horse; however, for the sake of simplifying the argument, I will presume that the reader is familiar with these terms and for those who are not familiar with them, I refer them to Emmon Bach, *An Introduction to Transformational Grammars* (New York: Holt, Rinehart & Winston, 1964) p. 69. I owe the information about Savin's experiments to Noam Chomsky (from my notes taken of his lectures at UCLA, summer, 1966).

of sentence structure gradually gets beyond the control of human memory in such a way that a sentence with more than two degrees of self-embedding is not readily usable in actual speech or even written language.[10] Generalizing on these psychological experiments and his theory of language, Chomsky assumes that the acceptability (usability) of a sentence in actual linguistic performance is attributable to the global properties of derivation, i.e., how linguistic rules are interrelated in its derivation to the organization of memory. In this vein, he hypothesizes that the organization of memory is such that it finds it strenuous to exercise control over a (matrix) sentence when an (insert) sentence of a similar structure is inserted in it. The strain on memory inscreases rapidly with the increase in the number of times that the sentences of a similar structure are embedded in the matrix sentence. In other words, the organization of memory is such that it can execute one simple operation at a time easily, i.e., scan such a simple sentence as "The boy is sick" easily at a time; but that it cannot execute simultaneously two or more similar operations so easily, i.e., cannot at one time scan the matrix sentence and insert sentences as in "The mascot that the soldiers that the Viet Cong that fled fought brought got stolen". The hypothesis that memory can execute very easily one operation at a time also explains the fact that branching structures, either left or right, are easily understood, memorized, and used in actual linguistic performance, because all that memory has to do in scanning branching structures is to execute its operations one by one either from left to right or from right to left.[11] Such branching structures as "men, women, boys and girls" are fairly easy for the memory to handle. The rules of transformational grammar such as self-embedding, branching, etc., and their correlation to the hypothesis about the organization of memory seem to provide a principled explanation for some aspects of our linguistic performance as discussed above. If so, transformational model as the theory of linguistic competence also provides a systematic basis for the fruitful study of linguistic performance.

It is now clear that transformationalists' concern with linguistic competence does not entail total disregard for the study of linguistic performance and that criticisms of transformational

[10] Self-embedding refers to the operation of a rule in transformational grammar, which inserts a structure into another structure, where, both structures are similar, e.g., inserion of "The boy killed the dog" into "The boy cut the class", as in "The boy that killed the dog cut the class". For a first hand report on the Miller and Isard experiment, see George A. Miller, "The Psycholinguists on the New Scientists of Language".

[11] A branching rule refers to a rule in transformational grammar, that strings together one after another cerain elements as in "rich, good-natured, handsome, and promising young man". For further discussion on transformational grammar as it relates to a theory of linguistic performance, see Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge: MIT Press, 1965).

model for neglecting linguistic performance are misplaced. On the contrary, some of the most fruitful insights into the theory of linguistic performance have been gained within the framework of transformational model.

4. Rule-Governed Creativity versus Simple Analogy

As soon as we accept as an empirical fact our unbounded ability in the production and interpretation of sentences, which we should for the reasons discussed, we face the problem of accounting for this infinite ability of finite human beings. We have argued that the only plausible account for this is to assume that human beings master or internalize a finite set of rules, the application of which gives rise to the infinite creativity in their use of language. In other words, rule-governed creativity is the only explanation for our linguistic competence.

This explanation of creativity in language use sharply contrasts with the Bloomfieldian or neo-Bloomfieldian view that analogical creation, association or inductive generalization is responsible for our ability to innovate in the use of language. Do we derive "His father is resembled by the boy", which is clearly incorrect, from the correct "The boy resembles his father" on the analogy of the derivation of "His father is respected by the boy" from "The boy respects his father"? The view of Bloomfield and his followers[12] fails to give a full and accurate account of our unbounded ability in language use, because 1) the examples like the above which show analogical creation to lead to absurd consequences are not confined to a small fraction of the language, 2) for analogy or equivalently inductive generalization to work at all for the entire language, a person should have encountered an infinite number of sentences or sentence types required for analogical operation during the short period of his life when he learned his language as a child (which possibility is ruled out because of the limited nature of the sample of speech, on the basis of which he learned language), and 3) the observable phonetic features of many sentences are so impoverished semantically that a straightforward association between the semantic elements and the phonetic features of sentences is often impossible without which analogical creation or inductive generalization cannot work.[13]

In the light of the many difficulties and inadequacies of the Bloomfieldian view of linguistic

---

[12] For discussions of Bloomfieldian type of analogical creation, see Leonard Bloomfield, *Language* (London: George Allen & Unwin, Ltd., 1934) Chapter 23; Charles F. Hockett, *A Course in Modern Linguistics* (New York: The Macmillan Company, 1958) pp. 425—438.

[13] For a discussion of rule-governed creativity versus analogical creation, see Noam Chomsky, *Cartesian Linguistics* (New York: Harper & Row, 1966) pp. 3—31.

innovation on the one hand and of the plausibility and empirical justifiability on the other of the transformational view in this regard, it seems that the latter view should be accepted as closer to linguistic facts known to us.

5. Underlying and Superficial Structures[14]

In some sentences, it is rather easy to associate the observable phonetic features with the meaning elements of the sentences. We can, for example, easily pair certain words with certain meanings they carry in such sentences as "John hit the ball". However, in many other sentences, such a straightforward association is not possible because there are no phonetically observable features in the sentences, to which certain obviously felt semantic constituents can be paired.

For instance, native speakers of English intuitively feel that an imperative sentence is addressed to the second person and refers to an action desired to be taken by the addressee sometime in the future. Quite frequently, however, imperatives do not contain any one of the various future tense markers and/or the second person subject "you". Take, for illustration, "Help yourself". Neither a future tense marker nor the second person subject is to be observed. Consider, however, that a reflexive pronoun, used as an object, always agrees with the subject in person, gender, number, and reference. This leads to the conclusion that the subject deleted in "Help yourself" sholud be "you". This logical conclusion is supported further by the consistent occurrence of "you" in such syntactic paraphrases of "Help yourself" as "You will help yourself", "You help yourself","Help yourself, will you", etc. The contention that an imperative refers to a future action is also supported formally by the observation of the future-marking "will" in two of the three paraphrases of "Help yourself" enumerated above.[15]

The above line of reasoning shows that "you" and "will" are always associated with an imperative whether or not they are realized phonologically. To provide a unique formal basis for the (semantic) interpretation of this range of facts, we might posit an abstract syntactic structure of the form "you+will+verb phrase+X+imp" for imperative sentences. Such an abstract structure as this is called the underlying structure and is designed to provide all information necessary for semantic interpretation. From this underlying structure will derive such superficially observable forms as the various forms of the imperative cited above by deletion

---

[14] For the most clear-cut discussion of the distinction between underlying and superficial structures, see Paul M. Postal, "Underlying and Superficial Linguistic Structures", *Harvard Educational Review*, Vol. 34, 1964. For a historical survey of the same topic, see Noam Chomsky, *Cartesian Linguistics*.

[15] The argument given here is adapted from Paul M. Postal, "Underlying and Superficial Linguistic Structures".

and/or rearrangement operations. The syntactic structure in such a superficially observable form, though not precisely, is referred to as the superficial structure. The superficial structure is so designed as to provide all information necessary for phonological interpretation.

5. 1. Motivation for the Distinction between Underlying and Superficial Structures

The distinction between underlying and superficial structures is very well motivated in view of the following considerations.

5. 1. a. Sometimes there is need to assign more than one semantic interpretations to a single sentence due to two or more conflicting syntactic relations in it. The expression phonemically transcribable as /ðə+kiliŋ+əv+ðə+táygərz/ may mean either "the killing of the tiger's", "the killing of the tigers", etc. due to different underlying syntactic relations that the superficial identity conceals. In all such cases, we have to have different underlying structures corresponding to different syntactic relations, though there is only one and the same superficial form, if we are to have a systematic formal basis for their disambiguation.

5. 1. b. Sometimes there is need to assign radically different semantic interpretations to superficially similar sentences due to radically different grammatical relations that distinguish them. Superficially, the grammatical relation between "John" and "please" seems to be the same in both "John is easy to please" and "John is eager to please". But "John" is the object of "please" in the former sentence and the subject in the latter sentence. Such radical differences in underlying grammatical relations, hidden by superficial similarities, should be indicated in the underlying structures of the sentences for a systematic semantic interpretation of those sentences.

5. 1. c. There is need to assign identical meanings to superficially different sentences, among which syntactic paraphrase relations hold.[16] We have already suggested the need to posit one underlying structure for the superficially different forms of an imperative.

Evidence like the above, plentiful in human languages, forces any linguistic theory to make the fundamental distinction between the underlying and superficial structures, if it is to account adequately for the ways in which native speakers produce and understand infinitely many new sentences coherently. It is the failure to make this empirically well-motivated fundamental distinction between the two structures that causes all nontransformational theories of language to be *ad hoc* or inadequate. Nontransformational theories of language recognize one structure per sentence, that roughly corresponds to our superficial structure, which is extremely

[16] Syntactic paraphrase relations hold between sentences, which have the same meaning due to the same grammatical relations in them irrespective of their actually observable formal differences.

unrevealing as to the meaning of the sentence. Inadequacies such as this in pre-transformational linguistics motivated the transformational model.

5. 2. Phrase Markers and Structural Descriptions

We have argued that an adequate theory of language must make the distinction between underlying and superficial structures. The formal representations of these dual structures of sentences are called phrase markers of the sentences. Corresponding to the underlying and superficial structure distinction, phrase markers are of two kinds, i.e., underlying and superficial phrase markers. The description of syntactic structures by means of phrase makers is called the structural description.

To recapitulate, underlying phrase markers will bear information about meanings while superficial phrase markers bear information about the actual forms, i.e., pronunciations of sentences. The native speaker's unbounded linguistic ability is reconstructed as the mastery of a finite set of rules, which can generate an infinite number of sentences, uniquely assigning to them in the process of generation correct structural descriptions.

6. Transformational Grammar and Its Organization

Transformational grammar is a finite system of rules so constructed as to generate all and only the sentences of a language with their correct structural descriptions of the above nature so that the pairing of meaning to pronunciation in each of the generated sentences can be predicted in an explicit and formal manner.[17] Transformational grammar is formal in that its rules are stated in formally defined terms. It is explicit in that the mechanical application of the formally stated rules results in the precise specification of the structural descriptions of the sentences generated so that nothing whatsoever is left to the understanding interpretation of an intelligent reader with regard to rule application. Naturally one of the major methodological concerns is the formalization of the notion rule of grammar that is formal and explicit enough in the above sense.[18] The requirement that transformational grammar generate *all* sentences means that the generation of sentences should be complete in the sense that it should enumerate all the sentences in the infinite set of sentences that speaker-hearers tacitly

[17] To be able to produce and understand sentences means to be able to pair meanings to pronunciations in them so that the task of grammar construction is the construction of a model that performs this pairing in infinitely many sentences.

[18] For a discussion of the problem of explicitly specifying the notion "rule of grammar", see Noam Chomsky, "The Notion Rule of Grammar" J. A. Fodor and J.J. Katz (eds), *The Structure of Language: Readings in the Philosophy of Language* (Englewood Cliffs: Prentice-Hall, Inc., 1965).

know; the failure to generate *all* the sentences would mean that the transformational grammar is incomplete as the description of a language. The requirement that transformational grammar generate *only the* sentences means that it should generate only the grammatical sentences and no non-grammatical sentences; the failure to generate *only the* sentences would mean that the transformational grammar cannot reconstruct the ability of speaker-hearers to tell grammatical sentences from ungrammatical ones.[19] The requirement that the grammar should generate sentences with *correct structural descriptions* means that it should assign, for instance, different structural descriptions to superficially similar sentences "John is easy to please" and "John is eager to please" in such a way that "John" is the object of "please" in the former but the subject in the latter; the failure to assign *correct structural descriptions* to sentences it generates would mean that it cannot reconstruct the ability of speaker-hearers to interpret and produce sentences with unique semantic messages paired to pronunciations.

As such, transformational grammar consists of 1) the syntactic component, which generates infinitely many abstract structures of formatives (that underlie sentences) with their correct structural descriptions consisting of underlying and superficial phrase markers,[20] 2) the semantic component, which assigns meanings to the structures given as the output of the syntactic component on the basis of their underlying structures, and 3) the phonological component, which assigns pronunciations to the structures given as the output of the syntactic component on the basis of informaion contained in their superficial structures.

The syntactic component is the central component of grammar in the sense that it generates all and only the abstract structures of the infinite set of sentences of a language with correct structural descriptions, on the basis of which the latter two components operate. On the other hand, the latter two components are purely interpretative in the sense that they merely serve to interpret the abstract structures given as the output of the syntactic component semantically and phonologically, not on their own but on the basis of the syntactically given information in the form of structural descriptions. The role of the syntactic component with respect to the entire grammar is to mediate the pairing of meanings to pronunciations for all the sentences of the language.

---

[19] Grammatical sentences are generated by the application of the rules of grammar only, while ungrammatical ones are said to be derivatively generated by the violation of the rules of grammar at some point in the generation process. By "only the sentences" requirement, we do not mean that derivatively generated sentences are to be legislated against; on the contrary such sentences as "Colorless green ideas sleep furiously", though derivatively generated, may become a poetically effective line.

[20] The syntactic component consists of two subcomponents that generate two different types of phrase markers. The first subcomponent generates underlying phrase markers, on the basis of which the second subcomponent generates superficial phrase markers.

Incidentally, such an organization of grammar brings out very clearly the Saussurian dictum that the relation between the meaning and the pronunciation of a sentence or any other linguistic unit is arbitrary. For meaning and pronunciation of a sentence are obtained through the operation of two distinct components on two different structures of the sentence. In this way, there is only an indirect relation between the meaning and the pronunciation of a sentence.[21]

Traditional grammar was also concerned with the ability of men to produce and interpret any arbitrary sentence of their language. It made the correct distinction between underlying and superficial structures of sentences as in modern transformational grammar, and recognized the creative aspect of language use as the defining characteristic of language. It, however, failed to formalize precisely and explicitly rules of grammar and structural descriptions. Traditional grammar merely lists some examples to show its concern with the structural descriptions of sentences and the creative syntactic processes, listing exceptions and irregularities. It thus fails to go beyond listing and classification of some representative creative syntactic processes and structural descriptions. Consequently, it fails to account for all the sentences of a language, i.e., fails to be exhaustive With the rules of grammar unformalized, traditional grammar left too much to the interpretation of an intelligent reader, which is often haphazard. The lack of an adequate tool, i.e., a system of formal and explicit rules, accounts for traditional grammar's failure to be an adequate description of linguistic competence, which was its proposed goal. This shows that even a basically correct view of language, if not explicitly formalized, will be inadequate for the description of language. Thus the unformalized character of traditional grammar not only obscures its basic insights into the nature of language but also makes its claim to be a theory concerned with the creative aspect of language use rather doubtful.[22] This is an indirect argument for the explicit and formal specification of rules in tranformational grammar.

Modern descriptive linguistics commonly called taxonomic linguistics utterly fails to come to grips with the defining characteristic of language as a system of *rules*, considering quite erroneously that language is fully defined as a system of *elements* such as phonemes, morphmemes, IC's, etc. The proponents of taxonomic linguistics believe that the scientific description should be concerned only with "objectively" identifiable (observable) features of language, rejecting the mentalistic view that there are underlying creative processes beneath the surface of langu-

[21] This observation is due to J. J. Katz and Paul M. Postal, *An Integrated Theory of Linguistic Descriptions* (Cambridge: MIT Press, 1964) p. 2.

[22] Though defective as an unformalized grammar, traditional grammar has offered Chomsky many of the basic insights into the nature of language, which is incorporated in one way or another in his theory. For account of this, see Noam Chomsky, *Cartesian Linguistics* (New York: Harper & Row, 1966).

age.[23] Consequently, they paid scarcely any attention to the notion of rule of grammar, satisfied with the discovery of inventories of phonemes, morphemes, tagmemes, etc., for the languages analyzed. They have applied analytic-discriminative procedures to the observable features of a language in a limited amount of data, which is often erroneously considered large in the sense that the data contain all the relevant contrasts such as phonemic contrasts.

Modern taxonomic linguistics fails on two counts, among other things: it is 1) concerned with the highly unrevealing superficial forms of linguistic structures, and 2) it considers the description of a language complete when its segmentation-classification procedure establishes a system of elements and certain patterns, not realizing that the number of sentences is infinite and that only recourse to a system of rules can adequately describe language.

7. The Theory of Language and the Grammar of a Language

It is necessary to distinguish the theory of language and the grammar of a language. The theory of language is concerned with the common attributes of all languges, while the theory or grammar of a particular language is concerned with the peculiarities of that language.

Empirical investigations of languages reveal that there are linguistic features that are invariant from language to language. For instance, it is an empirical fact that we have to distinguish underlying from superficial structures in all languages thus far investigated. In addition to that, nouns, verbs, phonemes or distinctive features, regularities in tactics, such rules of grammar as grammatical transformations, etc., seem to appear in the description of all languages so far studied. Such invariant features recurring in the descriptions of all human languages, if fully enumerated and specified, would define the notion "the common attributes of human languages". These common attributes are called linguistic universals, the full specification and enumeration of which constitute the theory of the nature of human language.

Such a theory of language would be justified on two grounds: 1) it will offer a theoretical explanation as to why all human languages, despite their seeming diversities, have certain common features in their structures, and 2) individual linguistic descriptions, i. e., grammars of individual languages, will be immensely simplified if we enumerate and specify all the linguistic universals once in the theory and omit their mention in individual linguistic descriptions.[24] The degree of simplification thus achieved will be readily appreciated by consid-

[23] As a result, such absurdities as considering that the stem of "take" and "took" is /t-k/ and that /ey/, /u/, etc., are infixes, etc., all of which pose no problem if embedded in a system of rules, some of which will say take + past → /tuk/, etc.

[24] This argument is adapted from J.J. Katz, *The Philosophy of Language* (Harper & Row, 1966) pp. 109—110.

ering that the complication resulting from the specification and enumeration of linguistic universals in all individual linguistic descriptions would be the function of the number of linguistic universals and the number of individual languages described. Individual linguistic descriptions embedded in such a theory would be justified similarly: 1) they achieve significant generalizations to the effect that many of the feaures of individual languages are not their peculiarities but attributable to the general nature of language, and 2) they help achieve simplification in the manner described above. In the light of this, the theory of language is extremely well-motivated. The grammar of a language, not embedded in such a theory of language, will fail to achieve generalizations and simplifications and will be inferior to this extent.

It is important to note that the theory of language is an empirical, not *a priori* hypothesis about the nature of language, the linguistic universals being inferred from successful empirical descriptions of natural languages. Therefore, the justification given above is neither vacuous nor fortuitous. If we are to better understand the nature of language or a language, it is important at every stage of the empirical investigations of languages to look for linguistic universals and, if found, assign them to the theory of language.

It is convenient to classify linguistic universals as formal and substantive universals. Universal linguistc rules such as grammatical transformations, projection rules, etc., would qualify as formal universals; phonological distinctive features, semantic markers, syntactic markers, etc., which constitute the theoretical constructs out of which linguistic rules are formulated would qualify as substantive universals. The theory of language, to recapitulate, is the full enumeration and specification of formal and substantive universals.

If the theory of language is fully developed as such, grammars of individual languages would consist of systems of rules and constructs peculiar to them, the rest of their rules and coustructs being supplemented by the theory of formal and substantive universals. This is not to say that transformational model has at present reached such a stage but to say that such a theory construction is its ultimate goal.

8. Language Acquisition Device and the Construction of the Theory of Language

As repeatedly emphasized, the knowledge of language comes from the mastery of a system of rules, that enables us to produce and understand an infinitude of sentences. The mastery of this system of rules, which is by no means simple, takes place in the matter of a few years in our childhood on the basis of a sample of speech, technically called primary linguistic data. The primary linguistic data are 1) degenerate in the sense that they necessarily contain ill-

formed (i.e., ungrammatical) sentences as well as well-formed ones, 2) limited in the sense that they represent just a fraction of the infinite set of sentences of a language and also that they are presented in the matter of a few years, and 3) scattered in the sense that the sentences in the primary linguistic data are not organized in sequences as in language textbooks and that the primary linguistic data for one language acquisition situation necessarily differ from those for another. What is so surprisping is that all children in the same speech community acquire the same system of rules in about the same time despite the nature of the primary linguistic data discussed above. This, plus other considerations such as that 1)only human children learn language regardless of their intelligence without any meticulous training while the most intelligent chimpanzees or dolphins cannot even with the most meticulous training, 2) children are not predisposed to learn one particular language as against other languages, i.e., they can learn any language of the community where they are raised, and 3) there are commn attributes of languages known as linguistic universals, forces us to conclude that human children are endowed with a common species-specific device or capacity to acquire language. The human children seem to be equipped with this innately given device, which takes the primary linguistic data as its input and produces as its output the grammar of their language, i.e., a finite system of rules. This gives rise to their amazing ability to produce and understand an infinite set of sentences in a relatively short time.

What is the internal structure of this device, which enables the pairing of the primary linguistic data of the nature discussed earlier to a uniform grammar with such rapidity? At present, we can not give a precise answer to this question. However, on the basis of the nature of the input and that of the output we can advance a plausible hypothesis concerning the internal make-up of this device.[25] To produce a uniform output, i.e., the same grammar for all children in the same speech community on the basis of the degenerate, limited, and scattered input, the internal structure of the device must be highly complex and organized. Further, the structure of the device seems to be the same for all children, because 1) all children can learn any human language regardless of their intelligence in approximately the same time, and 2) all children in one speech community acquire the same grammar despite the discussed nature of the input.

Along this line of reasoning, Chomsky hypothesizes that there must be the following minimal

---

[25] For a detailed discussion of this device and its internal structure, see Noam Chomky, *Aspects of the Theory of Syntax* (Cambridge: MIT Press, 1965) pp. 30—37.

constituent elements that make up the internal structure of the device if we are to explain language acquisition at all:[26]

1) a universal theory of phonetics, like Jakobsonian distinctive feature theory, that defines certain phonetic signals as possible human sentences,

2) a universal notion of structural descriptions, that recognizes how the phonetic signals, defined as possible human sentences, are paired tentatively to meanings according to their structures,

3) a universal notion of the rules of grammar, that gives rise to certain initial hypotheses (alternative transformational grammars) about the language being acquired on the basis of primary linguistic data represented in terms of 1) and 2),

4) given any one of the hypotheses (i. e., grammars), a unique and mechanical method of assigning structural descriptions to any arbitrary sentence of the language, that predicts the meanings and pronunciations of sentences beyond the data,

5) an evaluation measure, that selects one of the alternative grammars as the highest valued.

A child, equipped with the language acquisition device of the above degree of sophistication, will, upon presentation of primary linguistic data, tentatively recognize the pairing of phonetic signals to meanings in sentences by virtue of conditions 1) and 2) in the device. Then it will form initial hypotheses, i.e., alternative grammars, for the language, of which the data are just a sample by virtue of condition 3) of the device. Then he or she tests the compatibility of the alternative grammars with the language by predicting sentences beyond the data and judging if the predictions conflict with the way the sentences about which predictions are made are understood and produced by the native speakers of the language,by virture of condition 4).[27] If the predictions made by a certain alternative grammar conflict with further data of the language too much the child will discard this grammar as unworkable for the language at hand. If, on the other hand, the predictions made by other alternative grammars are fairly close to the way the native speakers of the language use the sentences, then the child will revise them so that their further predictions may become completely compatible with the language. This process of testing, discarding, and revising hypotheses, i. e., grammars, goes on and on until all the remaining hypotheses are compatible with the language. Then condition 5) of the device comes in to select from among the remaining grammars the one most highly valued grammar,

---

[26] For a slightly different list of the minimal constituent elements of the language acquistition device, see J.J. Katz, *The Philosphy of Language*, p. 269.

[27] A grammar is said to be compatible with a language if it predicts any arbitary sentence of the language by assigning correct structural description(s) to it.

this highest valued grammar being the maximally simple one in the sense that it is the simplest and yet capable of predicting the sentences of the language correctly just as well as the others eliminated at this final stage of language acquisition.[28]

In the preceding section, we defined the theory of language as the full enumeration and specification of linguistic universals, formal and substantive. As a matter of fact, the minimal constituent elements that make up the internal structure of the language acquisition device are hypothetically postulated on the basis of formal and substantive universals in their varying roles in defining the nature of language. The universals are considered to be the reflection of the universality of the internal structure of the device. In other words, the hypothesis about the structure of the device is an extrapolation from the system of linguistic universals that constitutes the theory of language, the former being considered the antecedent cause for the latter. Viewed in this way, the construction of a hypothesis about the internal make-up of the device is the construction of the theory of language, which will provide a rational explanation not only for language acquisition but for the nature of language in general. We can now explain that the human ability to use language, i. e., to produce and understand an infinite set of novel sentences of any language is acquired as the consequence of the structure of the inborn language acquisition device exposed to the primary linguistic data from that language.[29]

In contrast to this essentially rationalist hypothesis, there has been a heavily behaviourist-oriented hypothesis about language acquistition. This opposing hypothesis says that the only role of what we have called the language acquisition device is simple association and inductive generalization, its internal structure being essentially blank. That this hypothesis is wrong is shown by the fact that 1) it is impossible to have encountered in our finite life-span all the sentences required for association and inductive generalization to work,[30] 2) the sentences presented in the primary data are scattered in the sense that they are not organized for the optimal operation of association and inductive generalization, 3) many sentences are semantically too impoverished to allow association and inductive generalization to work non-arbitrarily, etc.

9. The Evaluation Measure

The evaluation measure is part of the empirical hypothesis concerning the internal

---

[23] For a detailed discussion of the evaluation measure, see the next section.

[29] The role of the primary linguistic data in language acquisition is to 1) set into operation the language acquisition device, and 2) determine the direction in which the device will develop.

[30] This is because the number of novel sentences is infiite, for the acquisition of which we must enounter an infinite number of sentences.

structure of the language acquisition device that pairs the primary linguistic data to a grammar or grammars of a language. It is an empirical fact that all concrete attempts to construct a grammar for a language resulted in many mutually inconsistent grammars, which are all compatible with the data. It may be assumed that children acquiring a language also construct various alternative grammars compatible with the data of the language. The problem then is which of the alternative grammars the children innately prefer, so that the amazing rapidity and uniformity of language acquisition in approximately the same period of time can be accounted for. Suppose that no internal basis is given for the selection of one of the alternative grammars in the structure of the device. Then all the alternative grammars will be equally favored and there would hardly be any explanation for the equally rapid and uniform acquisition of language for all children of the same speech community irrespective of their relative intelligence.[31] A particular proposed evaluation measure is a hypothesis concerning the innate predisposition of a child to favor one alternative grammar over all others, postulated on the basis of an empirical assumption about the nature of association between the primary linguistic data and the resulting grammars, which is mediated by the device.

Empirical investigations of the structures of languages might reveal that significant linguistic generalizations. i.e., underlying regularities of linguistic structures, can be expressed in a certain notational system such as the one used in standard literatures on TG but not in any other notational system. Then we might assume that the nature of language is such that a grammar utilizing that particular notational system only captures the maximum degree of underlying regularities, i.e., achieves the maximum degree of liguistic generalizations. The evaluation measure, as proposed by transformationalists, operates in terms of the expressibility of linguistically significant generalizations or underlying regularities utilizing the smallest number of symbols. It is assumed that children acquiring a language also look innately for the alternative grammar that achieves the highest degree of significant linguistic generalizations using the smallest number of symbols over all others that fail to do so.

As a concrete example, let us consider the rule of the form Aux→tense (modal) (perfect) (progressive). This rule by virtue of the parenthesis notation captures the underlying regularity in the English verbal auxiliary system by collapsing into one rule a set of rules such as 1) Aux→tense, 2) Aux→tense+modal, 3) Aux→tense+perfect, 4) Aux→tense+progressive,

---

[31] Some of the alternative grammars might be cumbersomely long so that if no evaluation measure operates to eliminate them there would naturally be no explanation for the equal rapidity and uniformity observed in the language acquisition of all children.

5) Aux→tense+modal+perfect, 6) Aux→tense+modal+progressive, 7) Aux→tense+modal +perfect+progressive, etc. The use of the parenthesis convention here enables us to capture probably the maximum degree of the underlying regularity there is to the English auxiliary system. On the other hand, the mere listing of various possible combinations of auxiliary elements, factually correct as it is, fails to capture the significant generalization that underlies the seeming "irregularities". The use of the parenthesis convention, if supported by evidence from all the languages, may be posited as a formal property of grammar and amounts to the claim that only the grammar using this convention achieves significant linguistic generalizations, which will therefore be chosen against other alternative grammars not using this convention. A child will, it is assumed, also prefer such a grammar. In connection with this, it is important to remark that strong conditions on the form of grammar are often motivated by the desire to achieve linguistically significant generalizations, in terms of which the evaluation measure operates.

It is important to note that the evaluation measure is internal to a theory of language that aims at an explanatory hypothesis about the nature of language like the transformational model, and as such does not apply to competing theories of language not concerned with an explanatory hypothesis. A linguistic description may be simple in the absolute sense merely because it is not concerned with the crucial requirement that it achieve significant linguistic generalizations. The phonemic notation is simple in this sense but does not seem to be fit to express certain important generalizations and underlying regularities, which are expressible in terms of feature notation. Therefore, it is totally mistaken to claim, as some linguists do, that we should prefer phonemic notation to feature notation for the (mistaken) reason that the former is easier to read and saves more space than the latter.

10. Justification of Grammars

A theory of language that enumerates and specifies all linguistic universals including the evaluation measure will be able to provide a principled basis for the selection of the highest valued grammar for each language, providing an internal justification and explanation for the chosen grammar. The basis for the selection is principled because it is based on the empirically inferred nature of the association between the data and grammars in all languages investigated. The explanation and justification are internal because they come from the empirically inferred explanatory hypothesis about the antecedent cause for the acquisition of the chosen grammar, i.e., from the hypothesis about the internal structure of the language acquisition device, that

predetermines the nature of the selected grammar. A theory of language providing such a principled basis for the selection of one of the varirous alternative grammars as the highest valued and the grammar chosen by such a theory are said to be explanatorily adequate. An explanatorily adequate theory would meet all the five conditions of the language acquisition device.

A theory of language that makes available for each language a number of grammars compatible with the language are said to be descriptively adequate. Each of the grammars made available by such a theory is also said to be descriptively adequate and is justified on the external grounds that it generates sentences with structural descriptions in a way that matches native intuition. A theory of language, which meets the first four conditions of the internal structure of the language acquisition device, would be descriptively adequate in this sense, because it makes available a number of alternative grammars for each language, which are all descriptively adequate, i.e., compatible with the language. However, such a theory and the grammars it makes available will not achieve the level of explanatory adequacy, because no internal explanation on a principled basis is given for the selection of one of the grammars as the highest valued.[32]

To recapitulate, a grammar that is descriptively adequate will enumerate all the sentences of a language with correct structural descriptions; a grammar that is explanatorily adequate will not only enumerate all the sentences of a language with correct structural descriptions but also will be given an internal explanation as to why this grammar is favored over all others.

Only when a grammar is descriptively adequate, it begins to be of linguistic interest. Even a descriptively adequate grammar may leave unexpressed many things which are the defining attributes of language rather than the peculiarities of the particular language. Therefore, a descriptively adequate grammar is not fully adequate. Only an explanatorily adequate grammar is fully adequate in the sense that many of the features of the grammar, which are linguistic universals, will be explained as the universal attributes of language, that come from the universal internal structure of the language acquisition device.

Note that taxonomic linguistic models such as IC analysis achieve merely the level of some sort of observational adequacy since they are concerned with the superficial structures of

---

[32] A theory of language achieving merely descriptive adequacy will be concerned with the output of the language acquisition device, i.e., the grammars of languages, without any concern for the internal structure of the device; a theory of language achieving explanatory adequacy would be concerned with the nature of the device that pairs primary linguistic data to grammars; there is a still lower level of adequacy, observational adequacy, which is obtained when a grammar presents merely the facts observed in the input, the primary linguistic data, with no concern for the internal structure of the device or its output, i.e., grammar(s).

sentences; they emphasize the segmentation and classification of linguistic data on the basis of objectively observable superficial features in the data with no concern about correct structural descriptions or internal explanations about the superficially observable features in the data. On the other hand, the transformational model of linguistic description is concerned with the correct assignment of structural descriptions to the sentences generated by the model and with internal explanations in terms of an explanatory hypothesis about the internal structure of the language acquisition device.

Any theory, either linguistic or non-linguistic, becomes theoretically interesting only when it offers an explanation, hypothetically postulated or otherwise, about the facts observed in the field of its concern. If so, taxonomic models of linguistics are of no or little theoretical interest because they merely report facts observed in the surface structures of language without any concern for the deeper aspects of language, which will explain those facts observed.[33] If so, any person, who is interested in linguistics as a theory and not merely as a matter of observation, should reject taxonomic models and accept transformational model. This is not to deny the importance of taxonomic models' contribution to data gathering processes especially of unknown languages. It seems that enough linguistic data have been gathered and that it probably is the time to postulate theory to explain the collected data.

Selected Bibliography

Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambrige, Mass.: MIT Press, 1965.

_____. *Cartesian Linguistics*. New York: Harper & Row, 1966.

_____. *Current Issues in Linguistic Theory*. The Hague: Mouton & Co., 1964.

Katz, J.J. *The Philosophy of Language*. New York: Harper & Row, 1966.

Katz, J.J. and J.A. Fodor (eds.) *Tho Structue of Language: Readings in the Philosophy of Languge*. Englewood Cliffs, New Jersey: Prenetice-Hall, 1964.

Katz, J.J. and Paul M. Postal. *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass.: MIT Press, 1964.

Postal, Paul M. *Constituent Structure: A Study of the Contemporary Models of Syntactic Description*. The Hague:Mouton & Co., 1964.

[33] Taxonomic models do not even achieve observational adequacy in the strict sense of the word, because of their inherent failure to recognize the only way to account for the infinitude of novel sentences in a natural language is through the postulation of a finite set of rules of the nature discussed. That taxonomic models as such fail even to enumerate all the sentences has been demonstrated by Paul M. Postal, "Limitations of Phrase Structure Grammar", in J.A. Fodor and J.J. Katz, (eds.), *The Structure of Language*.